ACCURACY OF TRAFFIC COUNT DATA USED FOR CALIBRATION AND VALIDATION OF HIGHWAY MODELS

Reza Tolouei BEng MSc PhD MCIHT (corresponding author)
Principal Consultant

AECOM
AECOM House
63-77 Victoria Street
St Albans, AL1 3ER, United Kingdom
T +44 (0) 1727 53 5420
Reza.Tolouei@aecom.com

Javier Abellán MSc MEng
Senior Consultant

AECOM
AECOM House
63-77 Victoria Street
St Albans, Herts, AL1 3ER, United Kingdom
T +44-1727-53 5720
javier.abellancubria@aecom.com

Pablo Álvarez MSc(Eng) MSc MEng
Consultant
AECOM
AECOM House
63-77 Victoria Street
St Albans, Herts, AL1 3ER, United Kingdom
T +44-1727-53 5665
pablo.indave@aecom.com

1. INTRODUCTION

The sampling error in count data depends not only on the sample size, but also on the variation of individual observations during the sampling period. Furthermore, there are various other sources of uncertainty in the count data used for development of highway models (e.g. seasonal variations, uncertainty on vehicle splits, reduced sample size, etc.). There is lack of evidence on the overall accuracy of the count data used to develop highway models, taking into account various sources of error. It is therefore unclear how realistic the existing guidance (WebTAG/DMRB criteria) represent errors in the count data. This becomes increasingly important when emerging methods of trip matrix development (e.g. mobile phone data) are used instead of the conventional method of using Roadside Interview Data (RSI), where, unlike the latter, the expansion process is normally entirely independent from count data.

This study addresses these issues by quantifying sampling error in count data using statistical analysis techniques. Two multi-modal county-wide transport models, recently developed by AECOM, are used as case studies: the 'Leicester and Leicestershire Integrated Transport Model' (LLITM) and 'The Hertfordshire County

Model of Transport' (COMET). In both studies, highway matrices have been developed using mobile phone data and extensive count data have been used for calibration/validation of highway models.

The paper includes a discussion on how appropriate the existing WebTAG criteria are, given the expected range of errors in the count data, and whether alternative criteria should be considered in evaluating the performance of trip matrices.

The main objective was to investigate whether the magnitude of the sampling errors that exist in the data are consistent with the assumptions made within WebTAG. The paper also provides a summary of existing assumptions and criteria within UK, as well as examples from other countries.

2. EXISTING CRITERIA AND GUIDELINES

2.1. Evidence from the UK

The UK Department for Transport's WebTAG guidance provides advice on validation criteria and acceptability guidelines for highway assignment models. These are on the basis of assuming 95% confidence intervals of ±5% for two week Automatic Traffic Counts (ATCs), and ±10% to ±28% for one-day Manual Classified Counts (MCCs)[1]. These confidence intervals represent the sampling error in the count data.

When coming to the validation stage, the errors in the count data are not just caused by the sampling error, but also by other reasons such as the use of MCCs to split ATC data by vehicle type or the use of seasonality factors. The WebTAG validation criteria shown in the Table 1 are also an indication of the total errors in traffic count data at this stage.

Table 1 – WebTAG Individual Link Validation Criteria

Criteria	Description of Criteria	Acceptability Guideline
1	Individual flows within 100 veh/h of counts for flows less than 700 veh/h	> 85% of cases
	Individual flows within 15% of counts for flows from 700 to 2,700 veh/h	> 85% of cases
	Individual flows within 400 veh/h of counts for flows more than 2,700 veh/h	> 85% of cases
2	GEH < 5 for individual flows	> 85% of cases

For screenline flows, WebTAG criteria state that the difference between modelled flows and counts should be less than 5% of the counts for all or nearly all screenlines.[2]

It is understood that the WebTAG assumptions on errors in the count data, used as the basis of developing the criteria, are mainly sourced from The Design Manual for Roads and Bridges (DMRB). In particular, DMRB states the following (DMRB Volume 12 Section 2 Part 1, 3.2.20):

"Using the limited information available, the current best "working" estimate of the accuracy of measurement of the number of vehicles that passed an automatic traffic counter is that the 95% confidence interval of a count of longer than 12 hours duration is of the order of 5% of the total count". It also states that "Automatic traffic counts are usually the cheapest, simplest and most accurate and reliable form of count data available, […]"

2.2. Evidence from Elsewhere

A comparison is made between the criteria used by other countries in order to see the difference between different guidelines.

The US Federal Highways Administration (FHWA) validation guidelines argue that count data are often afforded more credence than they deserve and mentions that the variation of the count data should also be a concern in the development of the traffic count database.[3] It also states that setting validation standards for matching traffic counts can be a double-edged sword as it may encourage over-manipulation to meet the standards. FHWA explains the following:

"*a model validation that matches specified trip assignment standards within a reasonable range using valid modelling procedures is better than a model that matches observed volumes with a tighter tolerance using questionable modelling procedures."*

However, the FHWA does not provide standards, and the document highlights that matching or exceeding the guidance is not sufficient to determine the validity of a model. Instead, it reports some guidelines that have been used by various states and agencies. For example, the Michigan Department of Transportation (MDOT) has targets of 5 to 10 percent for screenlines, for the differences in observed and modelled flow volumes by screenline.

Figure 1 shows the maximum desirable deviation in total screenline volumes according to the observed screenline volume originally cited in Calibration and Adjustment of System Planning Models, produced by the FHWA in December 1990, and referenced in a number of other documents.

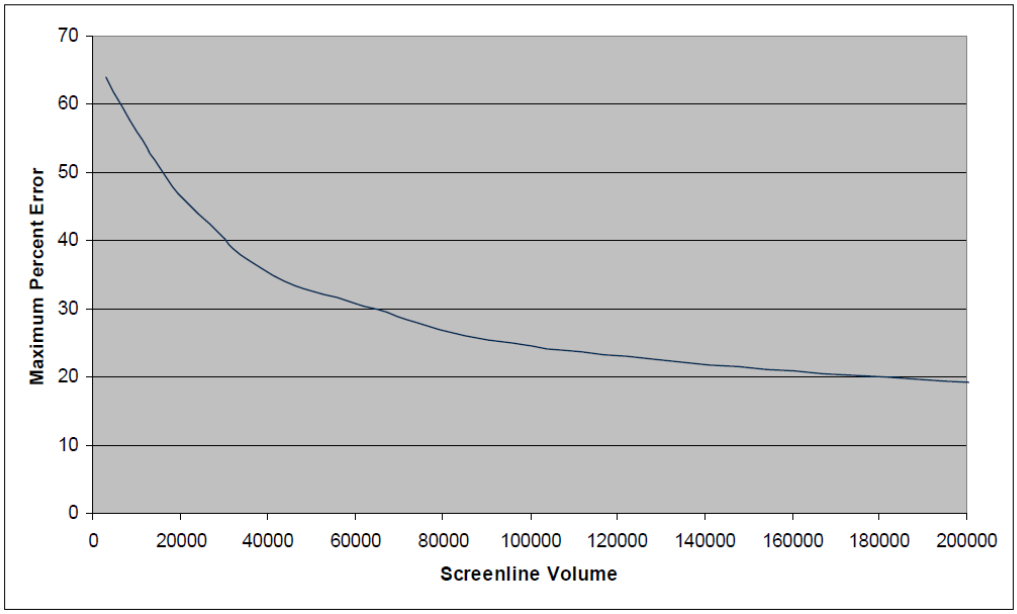

Figure 1 – Maximum Deviation, Calibration and Adjustment of System Planning Models, FHWA, December 1990

Table 2 shows the indicative guidelines used for the match between modelled and observed VMT for different states and aggregated by road type and area type.

Table 2 – Comparison of the Criteria used in Different States in the USA (Source: FHWA, 2010)

	Modeled Versus Observed VMT				
Stratification	**Ohio[a]**	**Florida[b]**		**Michigan[c]**	**FHWA-1990[c]**
Functional Class		**Acceptable**	**Preferable**		
Freeways/Expressways	±7%	±7%	±6%	±6%	±7%
Principal Arterials	±10%	±15%	±10%	±7%	±10%
Minor Arterials	±10%	±15%	±10%	±10%	±15%
Collectors	±15%	±25%	±20%	±20%	±20%
All Links		±5%	±2%		
Area Type					
CBD	±10%	±25%	±15%		
Fringe	±10%	±25%	±15%		
Urban	±10%	±25%	±15%		
Suburban	±10%	±25%	±15%		
Rural	±10%	±25%	±15%		

[a] Giaimo, Gregory, Travel Demand Forecasting Manual 1 – Traffic Assignment Procedures, Ohio Department of Transportation, Division of Planning, Office of Technical Services, August 2001.

[b] FSUTMS-Cube Framework Phase II, Model Calibration and Validation Standards: Model Validation Guidelines and Standards, prepared by Cambridge Systematics, Inc., for the Florida Department of Transportation Systems Planning Office, December 31, 2007, Table 3.9, page 3-16.

[c] The FHWA Travel Model Improvement Program Workshop over the Web, The Travel Model Development Series: Part I–Travel Model Estimation, prepared by Cambridge Systematics, Inc., June 9, 2009, Slide 11, http://tmip.fhwa.dot.gov/sites/default/files/presentation_8_with_notes.pdf, accessed November 29, 2009.

In Canada, there is not a specific modelling guideline to follow, and consultants are required to set the validation criteria for their models based on their previous experience or by referring to the US FHWA validation guideline commented in the previous paragraphs.

The Traffic Modelling Guidelines from the Australian Roads & Maritime Services (RMS)[4] mentions that traffic counts may have errors of up to 5% or 10%, depending on the count collection method (automatic or manual).

Table 3 shows the criteria used in Australia. Note that the guidance makes use of the Root Mean Square Error (RMSE), the GEH statistic and the R Square instead of a simple percentage or absolute value.

Table 3 – Criteria used in Australia, RMS.

Topic	Criteria
Link and Turn	Results to be tabulated in appendices and summarised in main report
	Tolerance limits for network-wide area:
	95 per cent of individual link volumes to have a GEH ≤ 5.0
	85 per cent of individual turn volumes to have a GEH ≤ 5.0
	All individual link and turn volumes should have GEH ≤ 10
	Plots of observed vs. modelled hourly flows required for all observations
	Plots to include lines showing GEH = 5 tolerance limits
	R^2 value to be included with plots and to be > 0.9
	Slope equation to be included with plots (intercept to be set to zero)
	All counts RMSE should be 30.0 or lower
Screenline or Cordon	Tolerance limits for network-wide area:
	Each directional screenline or cordon total to have GEH < 4.0

3. METHODOLOGY

The approach to estimate sampling error, represented as 95% confidence intervals, is set out in the following sections. These are described separately for individual counts and screenlines.

3.1. Sampling Error for Individual Traffic Counts

95% confidence interval of individual counts depends on the sample size and variation in the individual records. Assuming that the sample is taken from a population with unknown variance, the following formula can be used to calculate 95% confidence intervals; this can be used for count data of any sample size:

$$\mathrm{CI_{mean}} = \mathrm{m} \pm \mathrm{t_a}\left({}^{\mathrm{s}}/_{\sqrt{\mathrm{n}-1}}\right), \qquad (1)$$

where,

- *m* is the mean
- *CI_{mean}* is the 95% confidence interval for the mean
- *s* is the sample standard deviation,
- *n* is the sample size, and

t_a is taken from t distribution table with *df=n-1*.

The second term in the equation above is referred to as "level of accuracy", *E*. The percentage sampling error can then be estimated as *E/m*, or using the below equation:

$$\%\text{Error} = \text{CV} \times \left({}^{t_a}/_{\sqrt{n-1}} \right), \tag{2}$$

where *CV* is coefficient of variation (*s/m*). This equation shows that the two determinants of sampling error in the count data are sample size and variability of individual counts. Figure 2 shows how sampling error in the count data are related to sample size and variability in the counts, represented by coefficient of variation.

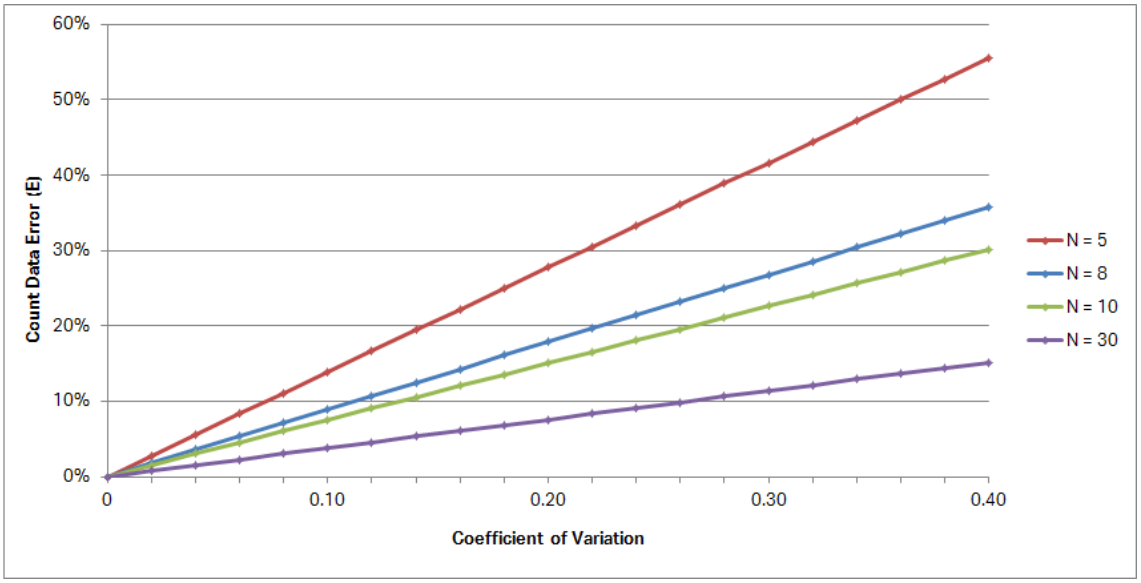

Figure 2 - Effect of Sample Size and Count Variability on Sampling Error in the Count Data

3.2. Confidence Interval for Screenlines

The variance of two random variables can be calculated as:

$$\text{Var}(x + y) = \text{Var}(x) + \text{Var}(y) + 2\text{Cov}(x, y) \tag{3}$$

If the two random variables are independent, the covariance term becomes zero. Traffic count surveys at each site are assumed to be independent; therefore, we can calculate the total error variance of the estimated traffic (E) associated with a specific screenline (SL) by summing the total error variances from each count site (1 to n).

$$\text{Var}(E_{SL}) = \text{Var}(E_1) + \text{Var}(E_2) + \cdots + \text{Var}(E_n)\ . \tag{4}$$

Given that the standard deviation of an estimate may be defined as the square root of the estimated variance of the quantity:

$$\mathrm{SD(t)} \equiv \sqrt{\mathrm{Var(t)}}. \tag{5}$$

the standard deviation of the errors for the screenline SL is computed as:

$$\mathrm{SD(E_{SL})} = \sqrt{\mathrm{SD(E_1)^2} + \cdots + \mathrm{SD(E_n)^2}} \ . \tag{6}$$

The standard deviation of the error in the sample mean for each individual count site is referred to as the standard error of the mean. For the total screenline counts:

$$\mathrm{SD(E_{SL})} = \mathrm{SE(M_{SL})}, \tag{7}$$

where M is the total screenline flow, which is sum of the mean flows at individual count sites.

The 95% confidence interval of the sum of the counts for the screenline SL can then be calculated as:

$$\mathrm{CI_{SL}} = \mathrm{M} \pm \mathrm{t_a SE(T_{SL})} \ . \tag{8}$$

t_a is taken from t distribution table with $df = \sum n - m$, where n is the sample size for each count site, and m is the total number of sites on the screenline.

4. CASE STUDIES

4.1. Background

COMET

In 2014, Hertfordshire County Council (HCC) commissioned AECOM to develop 'The Hertfordshire County Model of Transport' (referred to as "COMET"). COMET is a multimodal strategic model designed to forecast effects upon broad travel patterns and viability of corridors for investment. Figure 3 presents COMET simulation area, as well as Hertfordshire and Middle Layer Super Output Areas (MSOAs).

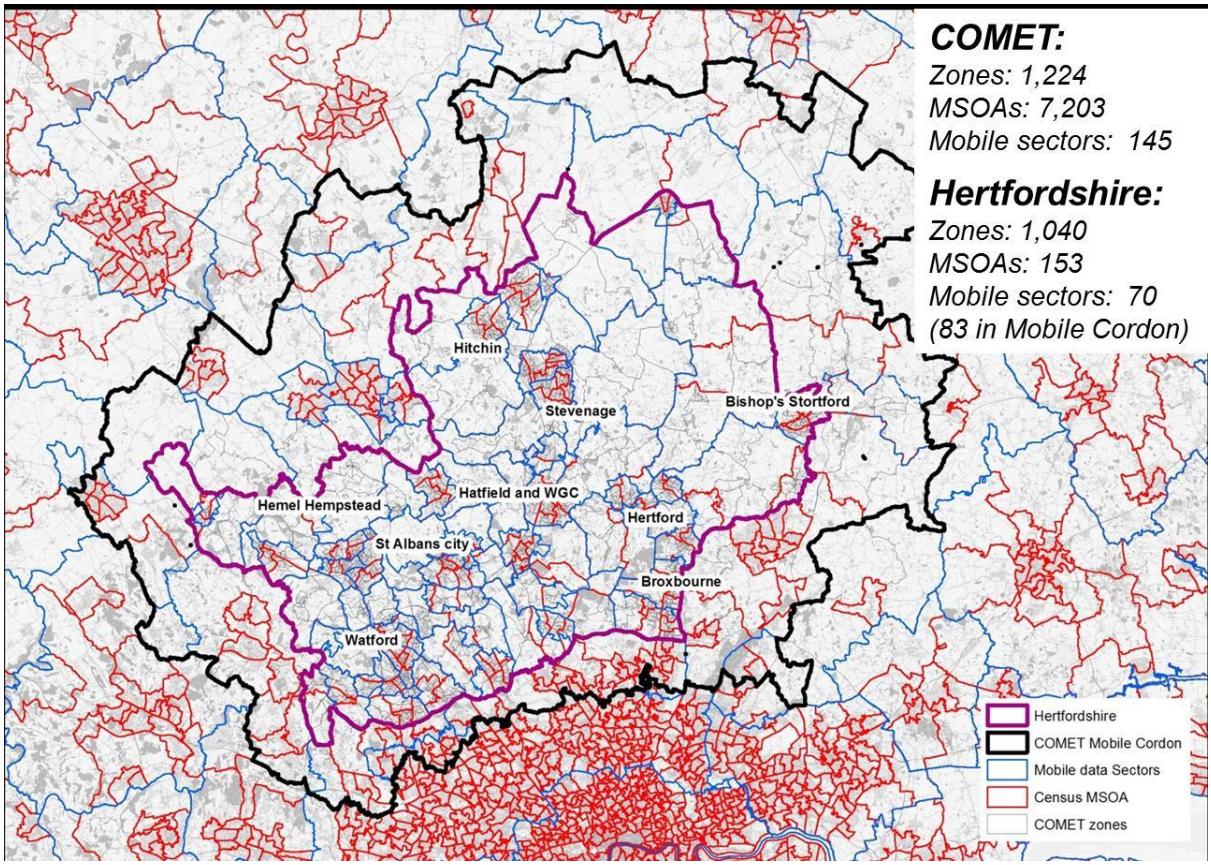

Figure 3 – Hertfordshire and COMET Zoning System

Various traffic data sources were considered for COMET development, including TRADS data, DfT Open Data, traffic data made available by Hertfordshire County Council (HCC) and traffic data from previous studies carried out by AECOM. However, these varied in age, quality and geographical distribution. More than 3,400 traffic data site locations were considered as part of the data review process. From these locations, 459 were analysed as part of the screenline and cordon definition process. The locations of these sites, together with defined cordons and screenlines, are shown in Figure 4 below.

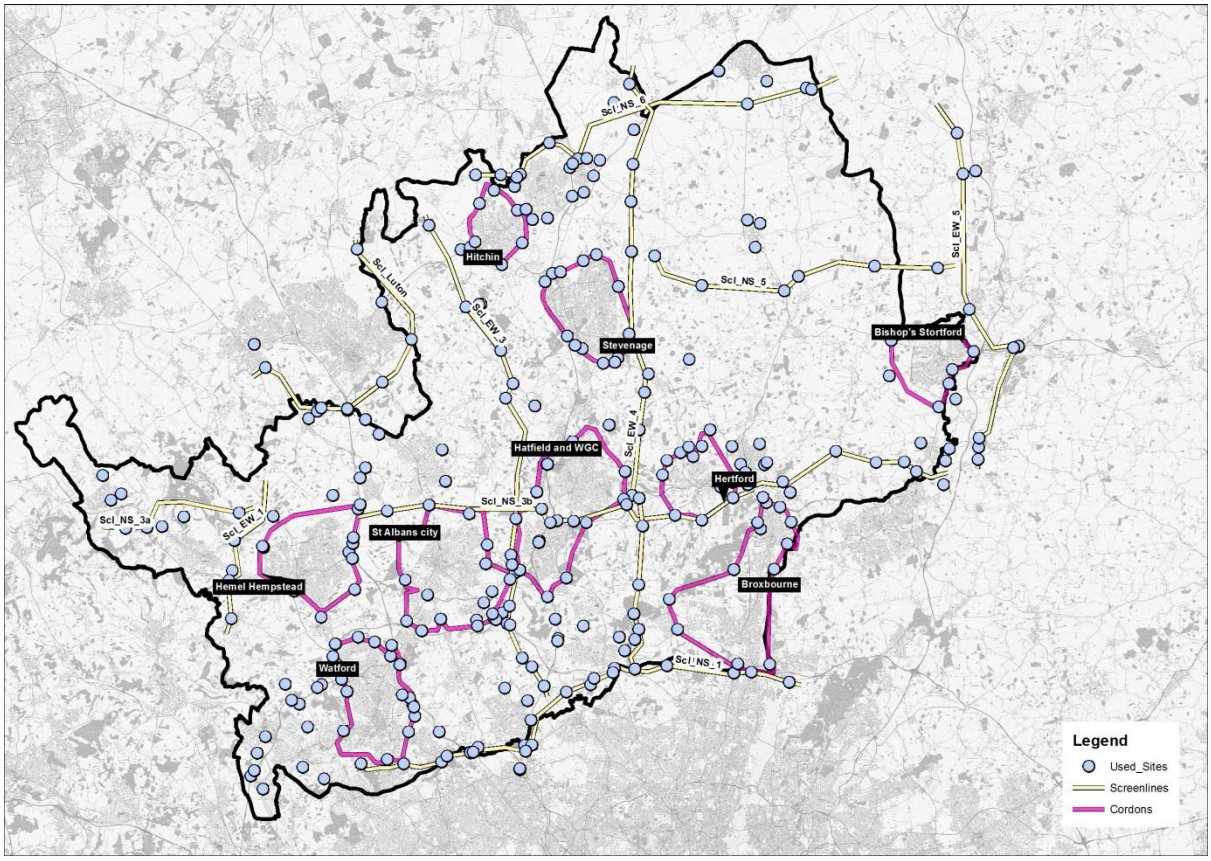

Figure 4 – Traffic Count Sites used in COMET Mode Development

In addition to the existing data, traffic surveys were carried out at 65 additional locations using Automatic Traffic Counts (ATCs). Manual Classified Counts (MCCs) were also carried out at 25 of these locations. Finally, traffic data from a total of 242 unique sites were used for the model calibration and validation. These data were grouped in 9 screenlines and 9 cordons around the main urban areas.

Data for these sites have been cleaned to ensure only the most recent available data are used, covering Monday to Thursday, excluding bank holidays and outliers. These data have also been processed to calculate average flows, standard deviations, and sampling errors for the AM and PM peak hours (0800 to 0900 and 1700 to 1800 respectively) and the IP average hour (1000 to 1600).

LLITM

The Leicester and Leicestershire Integrated Transport Model (LLITM) is a detailed multi-modal transport model of Leicester and Leicestershire in the UK, owned by Leicestershire County Council. It includes bus, rail, car, freight and walk and cycle modes of travel, as well as a detailed parking model in urban centres. AECOM is making a major update to LLITM to develop the LLITM 2014 Base model. As part of this, highway demand matrices are being developed using a combination of Roadside Interview data (RSIs) and mobile phone data from Telefonica.

The model is used for the development of business cases, assessing the impacts of new developments, underlining local policy and for strategic land-use and transport planning, and is one of the most sophisticated models of its type in the UK.

As part of the LLITM model development, Leicestershire County Council commissioned a significant programme of data collection. This included RSI data for about 155 sites (about 100 of which were surveyed in 2013/2014), local household survey data, local planning data, traffic count data from 663 sites, and local bus electronic ticket machine (ETM) data – See Figure 5 below.

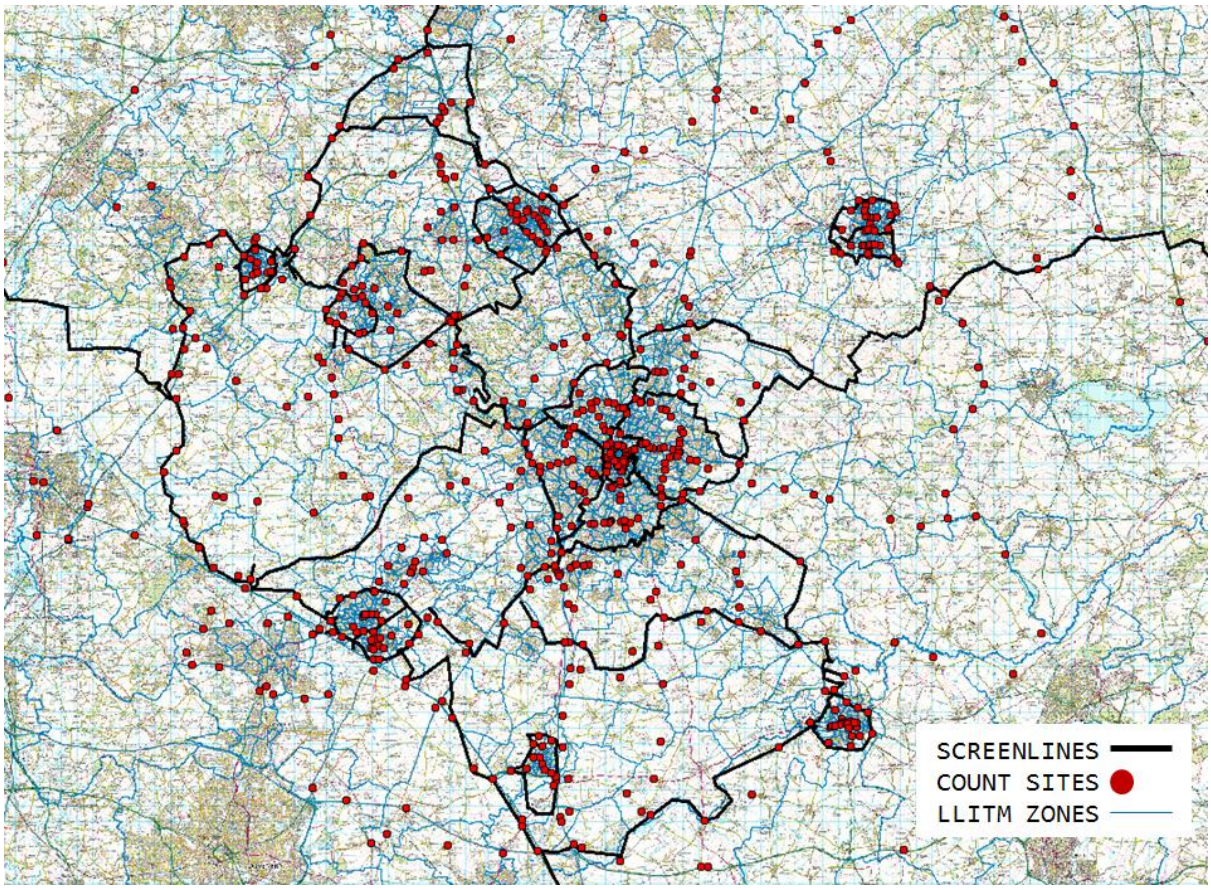

Figure 5 –Traffic Count Sites used in LLITM Mode Development

4.2. Sample Size

COMET

Figure 6 shows the distribution of sample size for all the available traffic counts that were processed for COMET. Almost 50% of the sites used in COMET have sample size of less than 8 days (less than two weeks). This is mainly due to the data cleaning process (removal of the outliers) and by the fact that, due to lack of data, some of these are processed data used in other studies which have made use of fewer observations. As mentioned in Section 2.1 above, WebTAG recommends using a minimum of two full weeks of ATC data where possible. The low sample size in some of these will have an impact on the sample error estimation since the sample error is heavily dependent on the sample size.

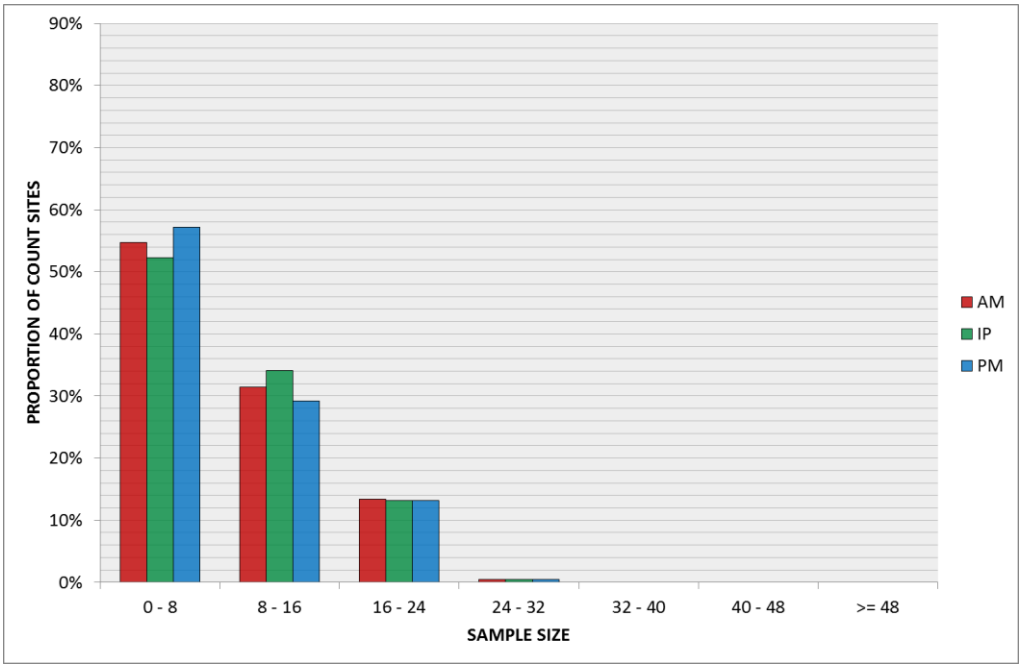

Figure 6 – Distribution of COMET sites by sample size

Table 4 shows the number of traffic count sites used for COMET and considered in this study as well as other descriptive statistics on sample size. It is important to note that only neutral days (Monday to Thursday) have been used[5]. Therefore, a sample size of 8 corresponds to a complete two-week ATC. COMET also uses data from sites with only one day of data, but these have been excluded from this analysis as is not possible to estimate the sampling errors for these sites. The average sample size for COMET is 8 days (two full weeks), with coefficient of variation being about 0.57.

Table 4 – Descriptive Statistics of Count Data Sample Size in COMET

Time Period	Sample size				
	No of sites (inc. direction)	Average	Minimum	Maximum	Standard Dev
AM	471	8	2	25	4.55
IP	469	8	3	25	4.51
PM	470	8	2	25	4.51

LLITM

Figure 7 below shows the distribution of sample size for all the available traffic counts that were processed for LLITM. Between 30% and 40% of the counts have a sample size lower than 8, which is mainly due to the cleaning process undertaken to remove outliers. The figure indicates that 40% to 50% of the counts have a sample size equivalent of 2 to 4 weeks, and 60% to 70% have a sample size higher than a two-week period.

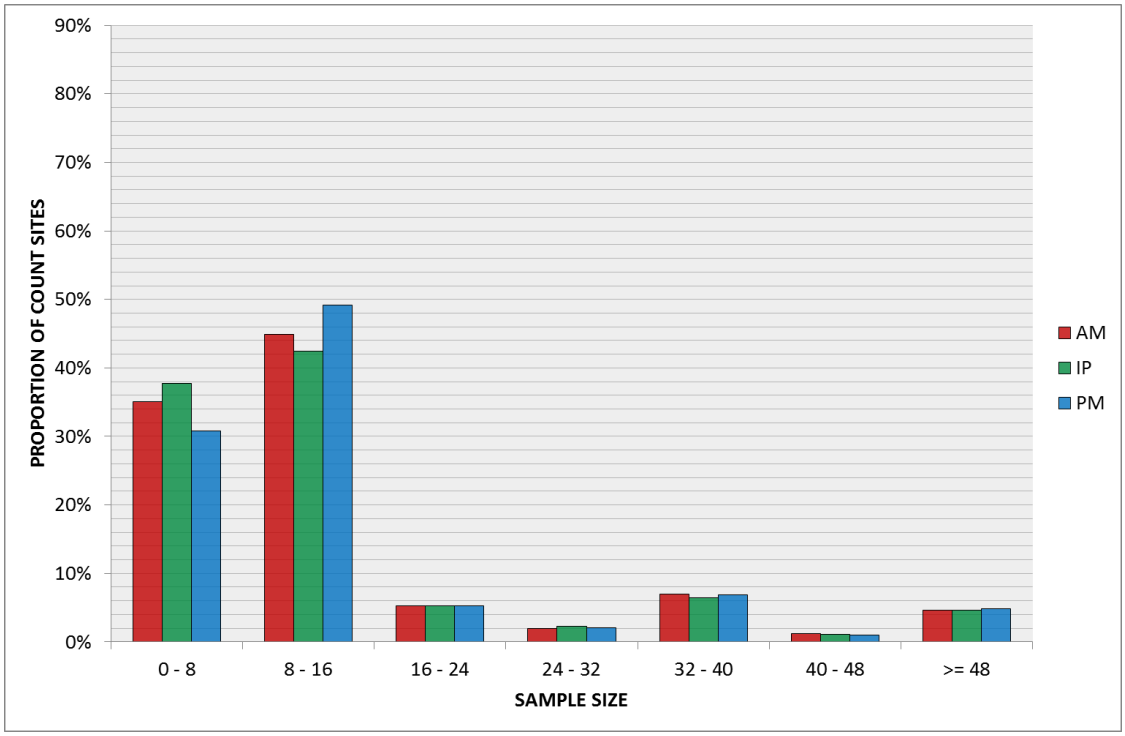

Figure 7 – Distribution of LLITM sites by sample size

Table 5 below shows the number of count sites used for LLITM development and considered for this study, as well as other descriptive statistics on sample size. The average sample size for LLITM is 15 days, with coefficient of variation being about 1.2.

Table 5 – Descriptive Statistics of Count Data Sample Size in LLITM

Time Period	Sample size				
	No of sites (inc. direction)	Average	Minimum	Maximum	Standard Dev
AM	1317	15	2	147	18
IP	1324	15	6	147	18
PM	1321	15	2	147	18

4.3. Sampling Error at Individual Traffic Count Level

The methodology set out in Section 3.1 was used to quantify sampling error, estimated as 95% confidence intervals, for COMET and LLITM count data by time period; the results are described in this section.

COMET

Figure 8 shows the distribution of sampling error for individual count sites in COMET, estimated for a subset of sites where the minimum sample size is equivalent of a two-week period (8 days). A significant proportion of sites, up to 30%, have a 95% confidence interval of 5% or higher.

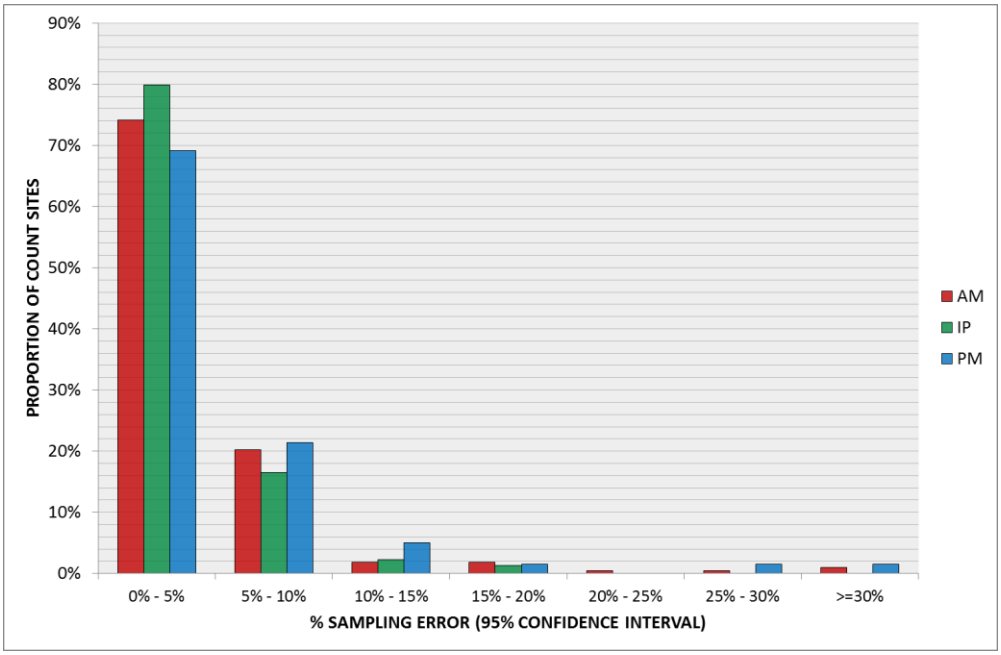

Figure 8 – Distribution of COMET sites by 95% Confidence Intervals

In terms of link flow validation criteria, WebTAG considers a greater tolerance. Table 6 shows proportion of sites where the sampling error is within this tolerance. As can be seen, most (but not all) of the sites studied have a sampling error within the tolerance allowed by WebTAG's link flow validation criteria. It is important to note, however, that sampling error is not the only source of error in the observed traffic counts; these are reflected in the difference between the assumed ±5% for 95% confidence interval and the criteria shown in Table 6.

Table 6 – Percentage of Counts in COMET where Sampling Error is within WebTAG's Validation Criteria

	AM	IP	PM	WebTAG	Equiv. percent.
Flow<700	98%	100%	97%	< 100 veh/h	14.3 %
700-2700	100%	100%	100%	15%	15%
>2700	100%	100%	96%	< 400 veh/h	14.8 %
ALL	99%	100%	98%		

LLITM

Similarly, sampling error was estimated for each traffic count site in LLITM. Figure 9 shows the distribution of sampling error, only taking into account sites in which the sample size is higher than a two-week period (8 days), as recommended in WebTAG. Depending on the time period, only between 50% and 70% of the counts have a 95% confidence interval within 5%, even where the minimum survey period is two week.

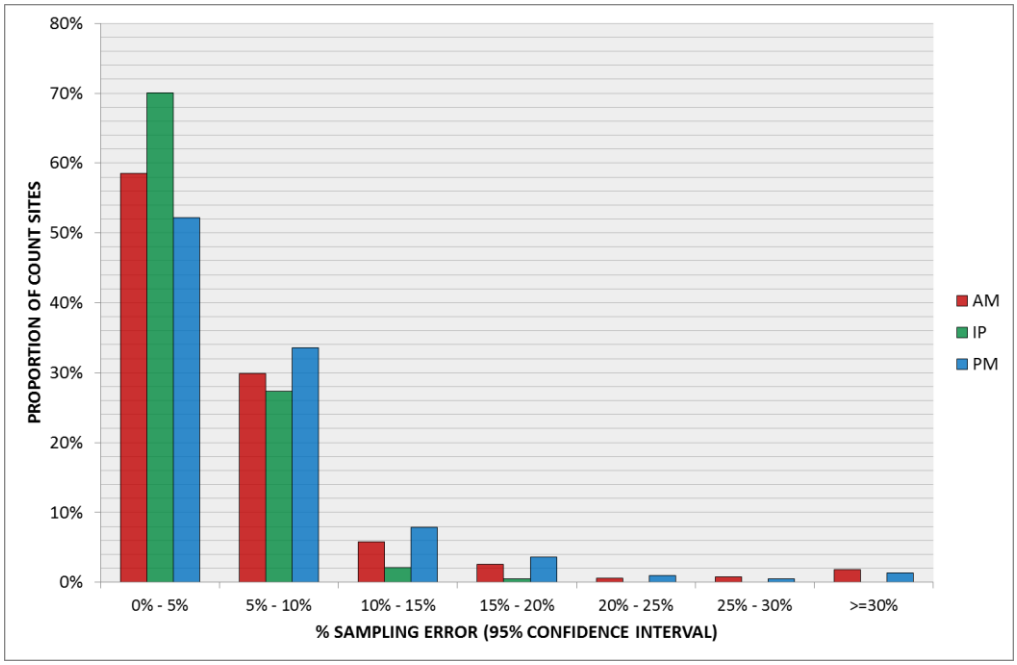

Figure 9 – Distribution of LLITM Sites by 95% Confidence Intervals

Comparing sampling errors with the range given as part of WebTAG's link flow validation criteria (Table 7), 96% to 99% of the counts have a sampling error within the allowed tolerance. However, as pointed before, other sources of errors that exist at the validation stage will contribute to the total error in the data; inclusion of these could therefore result in error ranges beyond the WebTAG values for a considerable number of count sites.

Table 7 – Percentage of Counts in LLITM where Sampling Error is within WebTAG's Validation Criteria

	AM	IP	PM	WebTAG	Equiv. percent.
Flow<700	97%	99%	96%	< 100 veh/h	14.3 %
700-2700	96%	99%	96%	15%	15%
>2700	96%	100%	98%	< 400 veh/h	14.8 %
ALL	96%	99%	96%		

4.4. 95% Confidence Interval sfor Screenlines and Cordons

The methodology set out in Section 3.2 was used to quantify sampling error, estimated as 95% confidence intervals, for the screenlines in COMET and LLITM models by time period; the results are described here.

Figure 10 and Table 8 suggest that in the COMET model, sampling error at screenline level is greater than ±5% in a significant number of sites (in up to 25% of a total of 18 screenlines and cordons, see Section 4.1 above). This implies that, at screenline level, sampling error of the count data are beyond WebTAG's recommended value of ±5% for about 17% and 25% of cordons and screenlines in AM peak and PM peak, respectively.

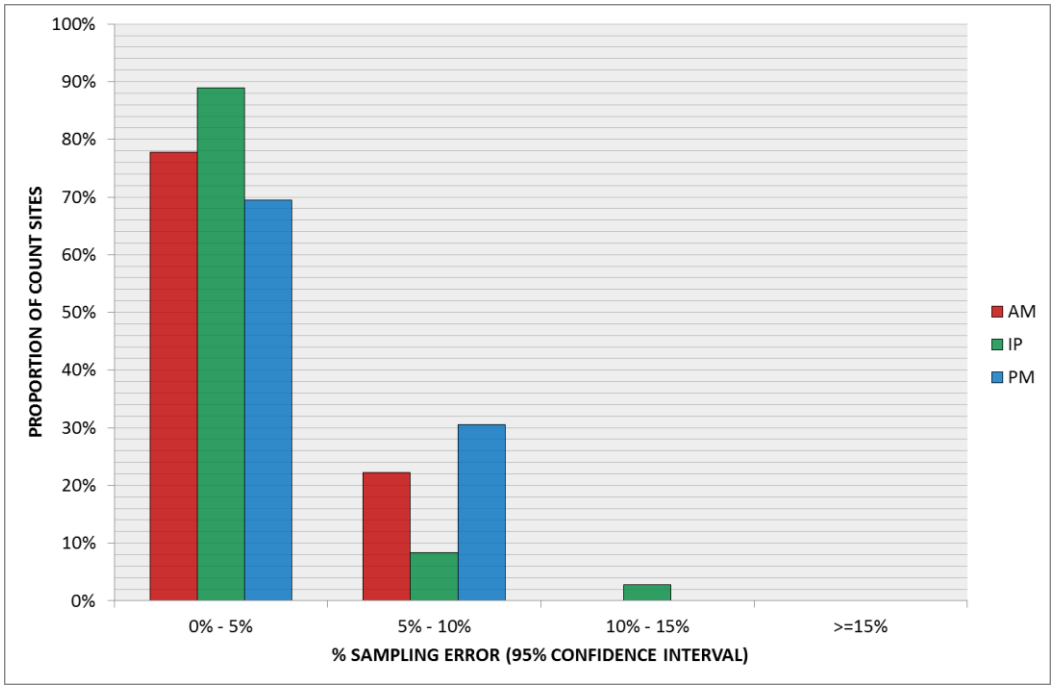

Figure 10 – Distribution of COMET Screenlines by Count Sampling Error

Table 8 – Percentage of Screenlines in COMET with Sampling Error within ±5%

	AM	IP	PM
ALL	78%	89%	69%

Figure 11 and Table 9 show a similar analysis for the LLITM model, where the distribution of 95% confidence intervals is shown by time period. As can be seen, almost all the screenlines have a confidence interval within 5% of the mean values.

A higher range of screenline errors in COMET compared with LLITM is partly due to a lower range of sample size (See Table 4 and Table 5 above), and partly due to the difference in the average length of screenlines and number of links in each screenline. COMET screenlines are longer and on average, include more sites than LLITM screenlines.

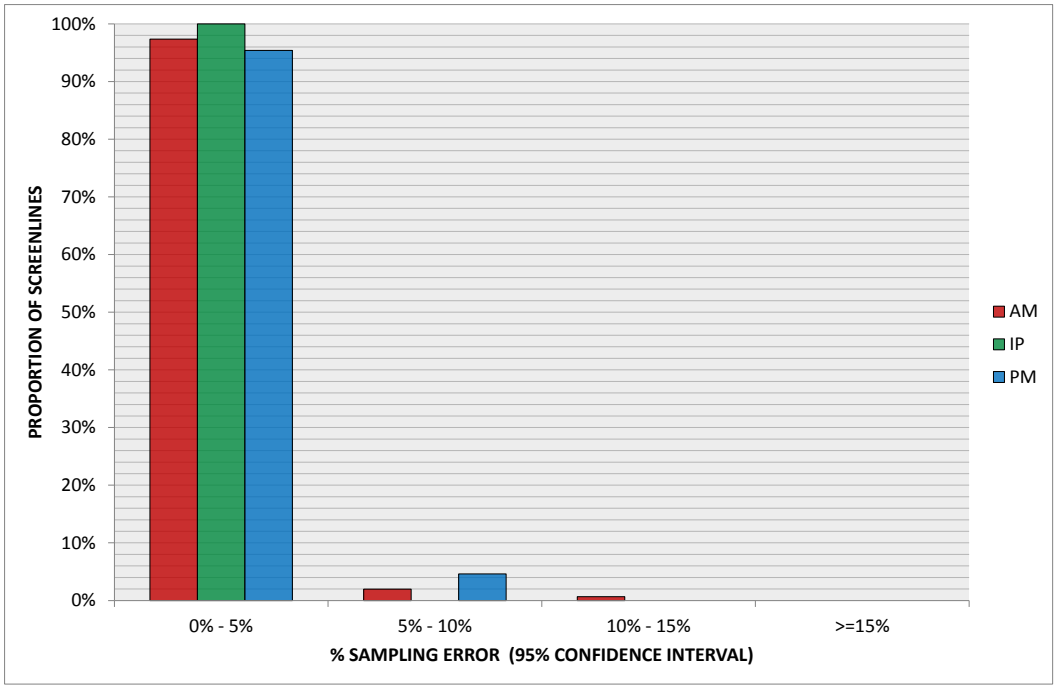

Figure 11 – Distribution of LLITM Screenlines by Count Sampling Error

Table 9 – Percentage of Screenlines in LLITM with Sampling Error within ±5%

	AM	IP	PM
ALL	97%	100%	95%

5. DISCUSSION AND CONCLUSIONS

Sampling errors in the count data used to validate two county-wide strategic transport models were quantified in this paper. The results were compared against the assumptions made in WebTAG, with a view to provide extra evidence on the consistency of these assumptions.

It was discussed in the paper that the current assumption on the 95% confidence interval of two-week ATC data within WebTAG is ±5%. It is understood that this is sourced from DMRB, as discussed in Section 2.1. It is not explicitly specified in WebTAG whether a two-week period referenced here includes certain weekdays only or all seven days of the week. However, WebTAG separately advises using survey data from Monday to Thursday, excluding weekends and Fridays. Therefore, in practice the above is usually interpreted as using ATC data which cover 8 days of traffic counts.

On the basis of the above, a subset of traffic count data, where the minimum sample size was 8 days, were used for each of the case studies and 95% confidence intervals were estimated for each count site, separately for each direction and time period. The results, reported and discussed in Section 4.3, showed a wide range of estimated 95% confidence intervals; these are represented as percent difference from the mean values. It was found that, depending on the case study and time

period, 10% to 30% of the counts (across all vehicle types) have a sampling error greater than ±5%, which is the assumption made by WebTAG, despite being taken from a sample of data with a minimum of two-week (i.e. 8 days) observations. This is because, as discussed in Section 3.1, sampling error depends not only on the sample size, but also on the day-to-day variability in traffic. This is an interesting finding which highlights a requirement for more focused research towards the appropriateness of existing WebTAG assumptions, to investigate whether any revisions on these are required.

In reality, most of the transport models, including COMET and LLITM, make use of count data where a considerable number of sites have a sample size less than the recommended 8 working days (about 30% and 50% of count sites for LLITM and COMET, respectively). This is partly explained by the cleaning process undertaken to identify and remove outliers in the data, resulting in a reduction in sample size. Therefore often a choice needs to be made between discarding any count data where the sample is low (due to various reasons), or using all the data which are available to the modellers. The latter is usually favoured to the former, justifying making the best use of all the information which is available, but accepting and acknowledging their limitations.

It could be argued therefore that a good practice in calibrating and validating highway models would be to estimate sampling error in the data where possible, and use this information when modelled flows are compared with the count data, allowing for more tolerance where sampling error is beyond the recommended values within WebTAG. It is important to note that the affected areas should therefore be clearly highlighted in the Local Model Validation Report, as model accuracy would be affected in these areas.

WebTAG's link flow validation criteria state that modelled flows should be within 15% of the count data (by vehicle type and time period). Whilst not clearly specified, the extra tolerance given here compared to the assumed ±5% sampling error, is probably to take into account other various errors in the count data such as application of temporal factors, and the uncertainty in vehicle type splits which are usually based on one-day MCC data. It is therefore arguable that, where 95% confidence intervals for a count location are greater than the assumed ±5%, this should also be translated to the link flow validation criteria to allow a wider tolerance in such cases. It is however recognised that, given lack of clarity in WebTAG on the source and the methodology used to identify the ±15% criteria for link flow validation, it would be difficult to quantify any extra allowance which should be given in such cases.

Finally, it should be noted that whilst the focus of this paper was on sampling errors in count data, there are other various sources of error related to the data which need to be taken into account when they are used to calibrate and validate modelled traffic

flows. The combination of the sampling error together with other sources of errors such as errors in MCC data or seasonality factors is expected to increase the overall error, as appropriately recognised by WebTAG. More research is therefore required to investigate and quantify various sources of error in traffic count data. In parallel, a clear understanding of errors in developed trip matrices is also necessary to inform any potential refinements in the model validation criteria, which is becoming more important with the increase in use of mobile phone data in building trip matrices.

NOTES

[1]This information is included in *WebTAG Unit M1.2* accessible from https://www.gov.uk/government/publications/webtag-tag-unit-m1-2-data-sources-and-surveys

[2] This information is included in *WebTAG Unit M3.1* accessible from https://www.gov.uk/government/publications/webtag-tag-unit-m3-1-highway-assignment-modelling

[3] The *Travel Model Validation and Reasonableness Checking Manual* is accessible fromhttps://www.fhwa.dot.gov/planning/tmip/publications/other_reports/validation_and_reasonableness_2010

[4] See Chapter 10 of the Traffic Modelling Guidelines, Roads and Maritime Services, NSW 2013.

[5] WebTAG, in its section 3.3.6 from unit *"M1.2. Data sources and survey",* indicates that a week from Monday to Thursday should be used.

BIBLIOGRAPHY

Department for Transport. (2013) *Transport analysis guidance: WebTAG*

Federal Highway Administration FHWA (1990) *Calibration and Adjustment of System Planning Models*

Federal Highway Administration FHWA (2010) *Travel Model Validation and Reasonableness Checking Manual*

Highways Agency (1998) *Design Manual for Roads and Bridges (DMRB)*

Roads and Maritime Services, NSW, (2013)*Traffic Modelling Guidelines, Roads and Maritime Services*